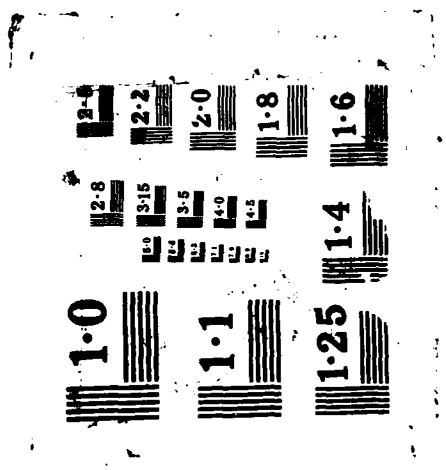
AD-A184 256   A HETEROSCEDASTIC HIERARCHICAL MODEL(U) CALIFORNIA UNIV   1/1
BERKELEY OPERATIONS RESEARCH CENTER W S JEWELL APR 87
ORC-87-11 AFORR-81-0122

UNCLASSIFIED   F/G 12/4   NL



OPERATIONS RESEARCH CENTER

A HETEROSCEDASTIC HIERARCHICAL MODEL**

by

William S. Jewell*

ORC 87-11 April 1987

UNIVERSITY OF CALIFORNIA

BERKELEY

11

A HETEROSCEDASTIC HIERARCHICAL MODEL[**]

by

William S. Jewell[*]

ORC 87-11 April 1987

[*]Department of Industrial Engineering and Operations Research, California
University of California, Berkeley, California 94720.

[**]This research has been partially supported by the Air Force Office of Scientific Research (AFSC), USAF, under Grant AFOSR-81-0122 with the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

AD-A184 25

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION: UNCLASSIFIED	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY: NA	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for Public Release; Distribution Unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE: NA	
4. PERFORMING ORGANIZATION REPORT NUMBER(S): 87-11	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION: University of Calif., Berkeley; 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: AFOSR/NM
6c. ADDRESS (City, State and ZIP Code): Engineering Systems Research Center, 3115 Etcheverry Hall, Univ. of Calif. Berk., Berkeley, California 94720	7b. ADDRESS (City, State and ZIP Code): Bldg. 410, Bolling AFB, DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR; 8b. OFFICE SYMBOL (If applicable): NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER
8c. ADDRESS (City, State and ZIP Code): Bldg. 410, Bolling AFB, DC	10. SOURCE OF FUNDING NOS.: PROGRAM ELEMENT NO. 6.1102F; PROJECT NO. 2304; TASK NO.; WORK UNIT NO. leave blnk

11. TITLE (Include Security Classification): A Heteroscedastic Hierarchical Model

12. PERSONAL AUTHOR(S): William S. Jewell

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
technical	FROM ____ TO ____		

16. SUPPLEMENTARY NOTATION
Keywords: Hierarchical models, collateral data, Bayesian prediction, heteroscedastic variances, credibility formulae.

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	see #16
XXX	XXXXXXXXXX	XX	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

see report

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT: UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION: UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL: Major Brian Woodruff	22b. TELEPHONE NUMBER: (202) 767-5027; 22c. OFFICE SYMBOL: AFOSR/NM

DD FORM 1473, 83 APR   EDITION OF 1 JAN 73 IS OBSOLETE.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

ABSTRACT

Hierarchical models are important in Bayesian prediction because they enable the use of collateral data from related risks with exchangeable parameters. The classical normal-normal-normal model with random means show clearly how the linear predictive mean for a single risk is improved by the availability of cohort data. However, this model has the disadvantage that the predictive density is *homoscedastic*, that is, the posterior variance depends only on the design (number of risks and number of samples). In most applications, one would assume that the variance also depended upon the data values.

One can, of course, change the variances at each level into random parameters, but this modifies the predictive mean formulae and leads to messy results in general. In the course of examining approximations to predictive variances, the author has found an extended normal model with variances that are quadratic in the data, and with the additional advantage that the linear mean formulae are unchanged.

Keywords: Hierarchical models, collateral data, Bayesian prediction, heteroscedastic variances, credibility formulae.

1. Introduction

Consider an *individual risk* (#1), characterized by an unknown *risk parameter*, $\tilde{\theta}_1$, from which n_1 i.i.d. *sample observations*, $\mathcal{D}_1=\{x_{1t} \; ; \; (t=1,2,\ldots n_1)\}$, are available; the problem is to predict a *future observation*, say $\tilde{w}_1 = \tilde{x}_{1,n_1+1}$. Given the *model density*, $p(x_{1t}|\theta_1)$, and *prior density*, $p(\theta_1)$, finding the *forecast density*, $p(w_1|\mathcal{D})$, is then a simple exercise in Bayes' law.

For a variety of simple likelihoods and priors (Jewell [1974] [1975a]), the forecast mean turns out to be a linear function of the data:

$$(1.1) \qquad \mathcal{E}\{\tilde{w}_1|\mathcal{D}_1\} = f_1(\mathcal{D}_1) = (1 - z_1)m + z_1(\Sigma x_{1t}/n_1) \; ,$$

where the mixing coefficient,

$$(1.2) \qquad z_1 = n_1/(n_1 + (e/d)) \; ,$$

is called the *credibility factor*, and the three required marginal moments are:

$$(1.3) \qquad m = \mathcal{E}\mathcal{E}\{\tilde{x}_{1t}|\tilde{\theta}_1\} \quad ; \quad e = \mathcal{E}\mathcal{V}\{\tilde{x}_{1t}|\tilde{\theta}_1\} \quad ; \quad d = \mathcal{V}\mathcal{E}\{\tilde{x}_{1t}|\tilde{\theta}_1\} \quad .$$

The *credibility formula*, $f_1(\mathcal{D}_1)$, has an obvious interpretation as a mixture of the prior mean with the sample mean, according to a learning curve, z_1, which tends towards unity with increasing sample size at a rate governed by a *time constant*, (e/d) . $f_1(\mathcal{D}_1)$ is a robust formula in the sense that it is also the best *linear least-squares fit* to the true $\mathcal{E}\{\tilde{w}_1|\mathcal{D}_1\}$ for arbitrary $p(x_{1t}|\theta_1)$ and $p(\theta_1)$ (Bühlmann [1967]) .

In many applications, the number of samples from risk #1 will be small, but there may be additional information available from *related risks*, that is, $(\tilde{x}_{it})$ characterized by a *different* risk parameter, $\tilde{\theta}_i$, but by the *same* form of model density, $p(x|\theta)$ $(i=2,3,\ldots r)$. For example, in insurance we may have a portfolio of risks which, a priori, are similar in nature, as measured by some risk classification scheme. Or, in health statistics we may have a cohort of apparently similar lives, with the same ages, heritage, diet, etc. Of course, if these risk parameters were independent, with the same density $p(\theta)$, then the *collateral data* would have no predictive value.

However, if we assume that $(\tilde{\theta}_1,\tilde{\theta}_2,\ldots,\tilde{\theta}_r)$ are *exchangeable*, we are able to keep the assumed similar nature of the related risks as well as to introduce correlation between the risks in a natural way; we do this by adding another unknown parameter, $\tilde{\varphi}$, which characterizes additional uncertainty at the portfolio or cohort level. In other words, the prior $p(\theta)$ is changed to a *conditional* prior $p(\theta|\varphi)$, and we assume a *hyperprior density*, $p(\varphi)$, is known, from which we can calculate the exchangeable *joint prior*:

$$p(\theta_1,\theta_2,\ldots,\theta_r) = \int \Pi p(\theta_i|\varphi)\ p(\varphi)\ d\varphi \ . \tag{1.4}$$

The resulting three-level structure is called a *hierarchical model*.

Unfortunately, the number of cases in which explicit analytic results can be obtained from hierarchical models is quite limited. As far as we know, the only case in which the predictive individual mean is linear in the individual and collateral data is in the normal-normal-normal linear model of Lindley & Smith [1972].

To simplify notation, we shall consider the important special case of this model:

$$(1.5)\qquad \begin{aligned} (\tilde{x}_{it}|\mu_i,\varphi) &\sim (\tilde{x}_{it}|\mu_i) \sim No(\mu_i;f) \quad ; \\ &\qquad\qquad (i=1,2,\ldots r) \\ (\tilde{\mu}_i|\varphi) &\sim No(\varphi;g) \quad ; \\ &\qquad\qquad (t=1,2,\ldots) \\ (\tilde{\varphi}) &\sim No(m;h) \end{aligned}$$

As before, we predict a future value of the risk #1, but now using the *total cohort data*, $\mathfrak{D} = \{x_{it} ; (i=1,2,\ldots r)(t=1,2,\ldots n_i)\}$. It will be apparent later that the predictive mean can now be written as the combination of two credibility-like forecasts:

$$(1.6)\qquad \begin{aligned} \mathcal{E}\{\tilde{w}_1|\mathfrak{D}\} &= f_1(\mathfrak{D}) = (1 - z_1)\ f_o(\mathfrak{D}) + z_1 y_1 \ ; \\ f_o(\mathfrak{D}) &= (1 - z_o)m + z_o Y_o \ . \end{aligned}$$

Here, the collateral data is organized into r+1 linear *sufficient statistics*:

$$(1.7)\qquad y_i = (\Sigma x_{it}/n_i) \ ; \ Y_o = (\Sigma z_i y_i/\Sigma z_j) \ ;$$

and r+1 credibility factors:

$$(1.8)\qquad z_i = \frac{n_i}{n_i + (f/g)} \ ; \ z_o = \frac{\Sigma z_i}{\Sigma z_j + (g/h)} \ ;$$

for each risk $(i=1,2,\ldots,r)$ and for the portfolio as a whole . $f_o(\mathfrak{D})$ turns out to be the predictive mean at the portfolio level, $\mathcal{E}\{\tilde{\varphi}|\mathfrak{D}\}$;

additional interpretations will be given below in Section 9. (1.6) is also the credibility formula (best linear predictive mean in the least-squares sense) for hierarchical models with arbitrary $p(x|\theta)$, $p(\theta|\varphi)$, and $p(\varphi)$ (Jewell [1975b]).

Thus, it would seem that the normal-normal-normal hierarchical model would be satisfactory for most situations in which collateral data is available, since, as we shall see below, the full-distributional predictive results are also easily obtained. Recently, however, the author and Hans Bühlmann have been examining credibility approximations for the *second* moments of $(\tilde{w}_1|\mathcal{D})$, in which various second-order statistics from the portfolio are used to augment a second-moment forecast using individual data (Jewell & Schnieper [1985], Jewell [1987], Jewell & Bühlmann [1987]). The least-squares analysis, while messy, is straightforward. However, (1.5) turns out *not* to be a useful test case for these approximate formulae, since all of the posterior second moments are *homoscedastic* in the data, by which we mean that they depend only upon the r+1 *sampling design parameters* (n_i, r), and not upon the actual data values. For instance, for all data $\mathcal{D}$:

$$(1.9) \qquad \mathcal{V}\{\tilde{\mu}_i|\mathcal{D}\} = g(1 - z_i) \quad ; \quad \mathcal{V}\{\tilde{\varphi}|\mathcal{D}\} = h(1 - z_o) \quad ;$$

and similarly for the predictive variance of the observables (see Section 4).

This is, we believe, a serious limitation of the normal-normal-normal model, since we would expect in a more general setting to be able to learn about any unknown variances from sufficient portfolio data. In what follows, we present our attempts to generalize (1.5) in order to obtain

heteroscedastic forecast formulae, while retaining the simplicity of the credible mean forecast (1.6). The resulting model turns out to be a generalization of the classical unknown mean and precision normal model for the non-hierarchical case. The formulation also clarifies the numerical integration that would be necessary to extend the generalization further.

2. A Generalized Hierarchical Model

To obtain heteroscedastical results from the normal hierarchical model, we must permit the variances at each level to be random quantities. Specifically, for risk i at time t, we assume:

$$(2.1)\qquad \begin{aligned} (\tilde{x}_{it}|\mu_i,\varphi,\omega,\tau,\eta) &\sim (\tilde{x}_{it}|\mu_i,\omega) \sim No(\mu_i;\omega^{-1}) \;; && (i=1,2,\dots,r) \\ (\tilde{\mu}_i|\varphi,\tau,\eta) &\sim (\tilde{\mu}_i|\varphi,\tau) \sim No(\varphi;\tau^{-1}) \;; && (t=1,2,\dots,) \\ (\tilde{\varphi}|\eta) &\sim \tilde{\varphi} \sim No(m;\eta^{-1}) \;. \end{aligned}$$

where $\tilde{\omega}$, $\tilde{\tau}$, and $\tilde{\eta}$ are new *random precisions*, corresponding to f^{-1}, g^{-1}, and h^{-1}, respectively. As before, the variable length *observed data*, $\mathcal{D}= \{x_{it}\}$, is to be used to estimate *all* of the *unknown parameters*, now denoted by $\tilde{\Theta} = (\tilde{\underline{\mu}}, \tilde{\varphi}; \tilde{\omega}, \tilde{\tau}, \tilde{\eta})$, plus make predictions of future $\tilde{x}_{it}$. $\tilde{\underline{\mu}}$ and $\tilde{\varphi}$ are the *conditional mean parameters*, as before. We let $\tilde{\Omega} = (\tilde{\omega}, \tilde{\tau}, \tilde{\eta})$ denote the group of unknown *precision parameters*; we assume that they are

statistically independent of the r+1 mean parameters. Temporarily, we assume that the precisions are governed by some known prior joint density, $p(\Omega)$.

Using the statistics y_i in (1.7), and setting $y_{ii} = (\Sigma x_{it}^2/n_i)$, we find the joint density of the data and the mean parameters, conditional on the precisions, as:

$$p(\mathcal{D},\underline{\mu},\varphi|\Omega) = (\tfrac{\omega}{2\pi})^{\Sigma n_i/2} (\tfrac{\tau}{2\pi})^{r/2} (\tfrac{\eta}{2\pi})^{1/2} \times$$

$$\exp\{-\omega \sum_i [n_i y_{ii} - 2\mu_i n_i y_i + n_i \mu_i^2]/2\} \times \tag{2.2}$$

$$\exp\{-\tau \sum_i [\mu_i - \varphi]^2/2 - \eta[\varphi - m]^2/2\}.$$

We then extract the conditional densities of the mean parameters in the usual tedious way.

At the individual risk level, we find that the conditional joint density of the mean parameters, $p(\underline{\mu}|\varphi,\Omega,D)$, is a product of *independent normal* densities, with means equal to:

$$\mathcal{E}\{\tilde{\mu}_i|\varphi,\Omega,\mathcal{D}) = f_i(\varphi,\Omega,\mathcal{D}) = [1-z_i(\omega,\tau)]\ \varphi + z_i(\omega,\tau)y_i \quad . \tag{2.3}$$

and variances:

$$\mathcal{V}\{\tilde{\mu}_i|\varphi,\Omega,\mathcal{D}\} = (\tau+\omega n_i)^{-1} = \frac{1}{\tau}[1 - z_i(\omega,\tau)], \tag{2.4}$$

for i=1,2,...,r. In parallel with the simpler model, we have defined *(conditional) individual credibilities*:

$$z_i(\varphi,\Omega) = z_i(\omega,\tau) = \left(\frac{\omega n_i}{\tau+\omega n_i} \right) \quad . \tag{2.5}$$

and (conditional) *individual credibility forecasts* $f_i(\varphi,\Omega,\mathfrak{D})$, for all risks. Note that neither the z_i nor the variances depend upon the data $\mathfrak{D}$ nor the highest-level precision parameter η. Thus, we are back to the homoscedastic case, as expected.

Removing $\underline{\mu}$ from (2.2), and again completing the square, we find that the conditional density of the cohort (portfolio) mean, $p(\varphi|\Omega,\mathfrak{D})$, is also normal, with a (conditional) *cohort credibility forecast*:

$$\mathcal{E}\{\tilde{\varphi}|\Omega,\mathfrak{D}\} = f_o(\Omega,\mathfrak{D}) = [1-z_o(\Omega)]\ m + z_o(\Omega)\ Y_o(\Omega), \tag{2.6}$$

where, as before, we define a (conditional) *cohort credibility factor*:

$$z_o(\Omega) = \frac{\tau\ \Sigma\ z_i(\omega,\tau)}{\eta + \tau\ \Sigma\ z_j(\omega,\tau)} \quad . \tag{2.7}$$

and the (conditional) *credibility-weighted average observation*:

$$Y_o(\Omega) = Y_o(\omega,\tau) = \frac{\Sigma\ z_i(\omega,\tau)\ y_i}{\Sigma\ z_j(\omega,\tau)}. \tag{2.8}$$

(We suppress the obvious dependence of Y_o upon $\mathfrak{D}$.) Note that z_o does not depend upon the data, and depends upon ω only through the (z_i). The variance of the conditional cohort mean parameter is:

$$\mathcal{V}\{\tilde{\varphi}|\Omega,\mathfrak{D}\} = \mathcal{V}\{\tilde{\varphi}|\Omega\} = (\eta + \tau\ \Sigma z_i(\omega,\tau))^{-1} = \frac{1}{\eta}\,[1-z_o(\Omega)]. \tag{2.9}$$

which is also homoscedastic, and depends upon ω only through the (z_i).

The remainder of (2.2) can now be matched with the still-general prior $p(\Omega)$ to give the conditional posterior joint density of the precision parameters as:

$$p(\Omega|\mathcal{D}) \propto p(\Omega)\ \omega^{\Sigma n_i/2}\ \{\prod_i[1-z_i(\omega,\gamma)]^{1/2}\}\{[1-z_o(\Omega)]^{1/2}\}\ \times$$

$$\exp\{-A(\Omega,\mathcal{D})/2\} \quad , \tag{2.10}$$

where the function A is:

$$A(\Omega,\mathcal{D}) = \omega\sum_i[n_i y_{ii} - n_i z_i(\omega,\gamma)y_i^2] + \eta\{m^2 - [1-z_o(\Omega)]^{-1}\ f_o^2(\Omega,\mathcal{D})\}. \tag{2.11a}$$

By expansion, and the use of the new definitions:

$$N = \Sigma\ n_i \quad ; \quad \bar{z}(\omega,\gamma) = N^{-1}\Sigma n_i z_i(\omega,\gamma) \ ; \tag{2.12}$$

$$Z(\omega,\gamma) = \ \Sigma z_i(\omega,\gamma)\ ;\ Y_{oo} = N^{-1}\Sigma n_i y_{ii}\ ;$$

the exponent can be put into several equivalent forms, such as:

$$A(\Omega,\mathcal{D}) = \omega[NY_{oo} - \Sigma n_i z_i(\omega,\gamma)y_i^2] + \eta z_o(\Omega)[m-Y_o(\omega,\gamma)]^2 \tag{2.11b}$$

$$- \gamma Z(\omega,\gamma)\ Y_o^2(\omega,\gamma) \quad ,$$

or

$$(2.11c) \qquad A(\Omega,\mathcal{D}) = N\omega \left\{ [Y_{oo}-Y_o^2(\omega,\gamma)] + N^{-1}\Sigma n_i z_i(\omega,\gamma)[Y_o^2(\omega,\gamma)-y_i^2] \right.$$

$$\left. + [1-\bar{z}(\omega,\gamma)][1-z_o(\Omega)][m-Y_o(\omega,\gamma)]^2 \right\} .$$

We apologize for this heavy notation, but it is important for the special models to follow so that we know explicitly where the various unknown precisions occur. In any case, it should be obvious from (2.10) and any of the (2.11) formulas that there is no "magic" prior $p(\Omega)$ that will give a tractable, closed-form posterior density, $p(\Omega|\mathcal{D})$!

3. Partially Unconditional Parameter Posteriors and Forecasts

Before progressing to various special cases, we first find several useful "marginal" densities, conditional only upon Ω and $\mathcal{D}$.

Noting that the exponent of (2.2) is quadratic in φ, and that $p(\underline{\mu}|\varphi,\Omega,\mathcal{D})$ is normal, we deduce that $p(\underline{\mu}|\Omega,\mathcal{D})$ must be *multinormal*, with moments defined as, say:

$$(3.1) \qquad \mathcal{E}\{\tilde{\underline{\mu}}|\Omega,\mathcal{D}\} = \underline{f}(\Omega,\mathcal{D}) \quad ; \quad \mathcal{V}\{\tilde{\underline{\mu}}|\Omega,\mathcal{D}\} = \Sigma(\Omega,\mathcal{D}) .$$

Unconditioning (2.3)(2.4) using (2.6)(2.9), we find, for the components of the mean vector:

$$(3.2) \quad [\underline{f}(\Omega,\mathcal{D})]_i = f_i(\Omega,\mathcal{D}) = [1-z_i(\omega,\gamma)]\, f_o(\Omega,\mathcal{D}) + z_i(\omega,\gamma)\, y_i \quad ;$$

and, for the components of the covariance matrix:

$$(3.3) \quad [\Sigma(\Omega,\mathscr{D})]_{i,j} = \sigma_{ij}(\Omega,\mathscr{D}) =$$

$$\begin{Bmatrix} \tau^{-1}[1-z_i(\omega,\tau)] + \eta^{-1}[1-z_i(\omega,\tau)]^2[1-z_o(\Omega)] & (i=j) \\ \eta^{-1}[1-z_i(\omega,\tau)][1-z_j(\omega,\tau)][1-z_o(\Omega)] & (i\neq j) \end{Bmatrix} .$$

with i,j = 1,2,...r. (Henceforth, we shall omit obvious ranges on indices.) Again, we see that even these partially unconditioned covariances are still homoscedastic.

It is also important to be able to find the *predictive densities* for future values of the individual risk variables. Let $\tilde{w}_{it} = \tilde{x}_{it}$ for *any* $t>n_i$. Since $\mathscr{E}\{\tilde{w}_{it}|\Theta\} = \mu_i$ and $\mathscr{V}\{\tilde{w}_{it}|\Theta\} = \omega^{-1}$, for all i and all future t, it follows from the above that $p(\underline{w}|\Omega,\mathscr{D})$ is also multinormal, with means:

$$(3.4) \qquad \mathscr{E}\{\tilde{w}_{it}|\Omega,\mathscr{D}\} = f_i(\Omega,\mathscr{D}) \quad ,$$

and covariances:

$$(3.5) \qquad \mathscr{C}\{\tilde{w}_{it};\tilde{w}_{ju}|\Omega,\mathscr{D}\} = \begin{Bmatrix} \omega^{-1} + \sigma_{ii}(\Omega,\mathscr{D}) & (i=j) & (t=u) \\ \sigma_{ii}(\Omega,\mathscr{D}) & (i=j) & (t\neq u) \\ \sigma_{ij}(\Omega,\mathscr{D}) & (i\neq j) & \end{Bmatrix} .$$

for all $t>n_i$ and $u>n_j$.

When considering the entire cohort, it may be useful to predict the *total cohort risk sum*, $\sum_i \tilde{w}_{it}$, or, more usually, the *cohort-average future risk*, $\tilde{w}_o = r^{-1}\sum_i \tilde{w}_{it}$ (for any future time $t > \max(n_i)$). Defining a *cohort-average credibility factor*,

$$(3.6) \qquad z_c(\Omega) = 1 - [1-z_o(\Omega)][1-r^{-1}\Sigma z_i(\omega,\gamma)] = [1+ \eta(r\gamma)^{-1}]\, z_o(\Omega) \quad .$$

we combine (3.2) and (2.6) to find:

$$(3.7) \qquad \mathcal{E}\{\tilde{w}_o|\Omega,\mathcal{D}\} = f_c(\Omega,\mathcal{D}) = [1-z_c(\Omega)]\, m + z_c(\Omega)\, Y_o(\omega,\gamma) \quad .$$

and:

$$(3.8) \qquad \mathcal{V}\{\tilde{w}_o|\Omega,\mathcal{D}\} = (r\omega)^{-1} + r^{-2}\, \Sigma\Sigma\sigma_{ij}(\Omega,\mathcal{D}) \quad .$$

Note that forecasting $\tilde{w}_o$ is not the same as estimating $\tilde{\varphi}$, since the former is from a finite, possibly biased, sample of individual risks that is fixed once and for all. The reader may easily find the covariance between successive future values of $\tilde{w}_o$.

4. The Constant Variance Model

As our first use of the above formulae, we note the obvious fact that setting:

$$(4.1) \qquad \omega^{-1} = f \quad ; \quad \tau^{-1} = g \quad ; \quad \eta^{-1} = h \quad ,$$

will give the full-distribution results for the simple hierarchical model described in Section 1. We get the simplification $Y_o = y_i$, and the various credibility factors become simply:

$$(4.2) \; z_i = \frac{n_i}{n_o+n_i} \quad ; \; z_o = \frac{\Sigma z_i}{r_o+\Sigma z_j} \quad ; \quad \bar{z} = r^{-1}\,\Sigma z_i \quad ; \quad z_c = [1+\frac{r_o}{r}]\,z_o \quad .$$

where $n_o = f/g$ and $r_o = g/h$. $p(\underline{\mu}|\mathfrak{D})$, $p(\varphi|\mathfrak{D})$, $p(\underline{w}|\mathfrak{D})$, and $p(w_o|\mathfrak{D})$ are multinormal, with means given by (2.3)(2.6) and (3.7), respectively. Interpretation of these results will be given below in Section 9.

As mentioned previously, all of the variances and covariances of these rv's are homoscedastic, in the sense that they depend upon the sampling design parameters (n_i, r), but not upon the actual data values in $\mathfrak{D}$. For instance, (3.3) simplifies to:

$$(4.3) \quad \mathscr{C}\{\tilde{\mu}_i ; \tilde{\mu}_j|\mathfrak{D}\} = \sigma_{ij}(\mathfrak{D}) = \left\{ \begin{array}{lr} g(1-z_i) + h(1-z_i)^2(1-z_o) & (i=j) \\ h(1-z_i)(1-z_j)(1-z_o) & (i\neq j) \end{array} \right\} .$$

5. A Special Non-Hierarchical Model

For our second special case, let us assume that knowledge about $\tilde{\varphi}$ is "tight", by letting $\tilde{\eta} \longrightarrow \infty$. This makes $\tilde{\varphi} \longrightarrow m$, almost surely, and effectively removes the hierarchical nature of the model. Thus,

$\tilde{\mu}_i \sim No(m;\tau^{-1})$, and the unknown parameters reduce to $\Theta = (\mu;\omega,\tau)$. Note also that, in the limit, $z_o \longrightarrow 0$, but that $(\eta z_o) \longrightarrow \tau\ \Sigma z_i(\omega,\tau)$. Furthermore, the $z_i(\omega,\tau)$ still depend upon both unknown precisions, so that the conditional joint density (2.10) is still rather complicated.

However, by examining the individual credibility factors (2.5), we see that they, in fact, depend only upon the *ratio* τ/ω of the remaining two unknown precisions! Therefore, if we *additionally* assume that $\tilde{\omega}$, say, has a prior density $p(\omega)$, and that the ratio is fixed at some positive value n_o:

$$\text{(5.1)} \qquad \tilde{\tau} = n_o\,\tilde{\omega} \quad ,$$

we obtain a great simplification, in that the $z_i = n_i/(n_o+n_i)$ are now *independent* of the precision parameters, as are the forecasts $f_i(\mathcal{D})$ in (3.2) (with f_o=m), and the sufficient statistic Y_o in (2.6) and (3.7). Thus, $p(\underline{\mu}|\omega,\mathcal{D})$ is multinormal with means $f_i(\mathcal{D})$ and simplified covariances:

$$\text{(5.2)} \qquad \mathcal{C}\{\tilde{\mu}_i;\tilde{\mu}_j|\omega,\mathcal{D}\} = \sigma_{ij}(\omega,\mathcal{D}) = \begin{Bmatrix} (\omega n_o)^{-1}(1-z_i) & (i=j) \\ 0 & (i\neq j) \end{Bmatrix} \quad .$$

In other words, the $(\tilde{\mu}_i|\omega,\mathcal{D})$ are conditionally independent. Similar results hold for the moments of the $(\tilde{w}_{it})$ and for $\tilde{w}_o$, where now $z_c = r^{-1}\Sigma z_i$.

More importantly, (2.10) now simplifies to:

$$\text{(5.3)} \qquad p(\omega|\mathcal{D}) \quad p(\omega)\ \omega^{N/2} \exp(-A(\omega,\mathcal{D})/2) \quad .$$

with the exponent:

$$(5.4)\quad A(\omega,\mathcal{D}) = \omega N \left\{ [Y_{oo} - Y_o^2] + (1-\bar{z})(1-z_o)(m-Y_o)^2 + N^{-1}\Sigma n_i z_i (Y_o^2 - y_i^2) \right\}$$
$$= \omega N\, B(\mathcal{D}) \text{ , say,}$$

now a *linear function* of ω! Thus, unconditioning w.r.t. ω can be carried out exactly, if we additionally assume that the *natural conjugate* gamma prior is used, that is:

$$(5.5)\qquad p(\omega) = \frac{\beta^{\alpha}\omega^{\alpha-1}\, e^{-\beta\omega}}{\Gamma(\alpha)} \text{ ,}$$

which we write as $\tilde{\omega} \sim Ga(\alpha,\beta)$.

It then follows that the posterior-to-data density is *closed under sampling*, and $(\tilde{\omega}|\mathcal{D}) \sim Ga(\alpha',\beta')$, with updated parameters:

$$(5.6)\qquad \alpha' = \alpha + N/2 \quad ; \quad \beta' = \beta + NB(\mathcal{D})/2 \quad .$$

In this way, the completely unconditional densities $p(\underline{u}|\mathcal{D})$, $p(\underline{w}|\mathcal{D})$, and $p(w_o|\mathcal{D})$ can be obtained as *Student-t* densities, and variances computed with the help of $\mathcal{E}\{\tilde{\omega}^{-1}|\mathcal{D}\} = \beta'/(\alpha' - 1)$.

We say that $\tilde{\underline{u}}$ has the *r-dimensional multivariate Student-t* density with ν degrees of freedom, mean vector $\underline{m}$, and precision matrix Ψ, written $\tilde{\underline{u}} \sim St_r(\nu;\underline{m};\Psi)$, if:

$$(5.7)\qquad p(\underline{u}) = \frac{\Gamma((\nu + r)/2)}{(\pi\nu)^{r/2}\,\Gamma(\nu/2)} \left(\frac{\nu}{\nu-2}\right) |\Psi| [1 + (\nu-2)^{-1}(\underline{u}-\underline{m})'\Psi(\underline{u}-\underline{m})]^{-(\nu+r)/2} .$$

with, of course, $\mathcal{E}\{\tilde{\underline{u}}\} = \underline{m}$, and $\mathcal{V}\{\tilde{\underline{u}}\} = \Psi^{-1}$. (r=1 is the ordinary Student-t density.) The proof that $p(\underline{\mu}|\mathcal{D})$ has this form follows from the fact that ω does not enter into the $f_i(\mathcal{D})$; it also follows that the $(\tilde{\mu}_i|\mathcal{D})$ are still independent but with density $St_1(\nu;m_i;\psi_{ii})$, where:

$$E\{\tilde{\mu}_i|\mathcal{D}\} = m_i = f_i(\mathcal{D}) \quad . \tag{5.8}$$

$$\mathcal{V}\{\tilde{\mu}_i|\mathcal{D}\} = \psi_{ii}^{-1} = (n_o+n_i)^{-1}\mathcal{E}\{\tilde{\omega}^{-1}|\mathcal{D}\} \quad . \tag{5.9}$$

A separate analysis shows there are $\nu = 2\alpha' = 2\alpha+N$ degrees of freedom. Similarly, $(\tilde{w}_{it}|\mathcal{D}) \sim St_1\left[(2\alpha+N)\ ;\ f_i(\mathcal{D})\ ;\ [1+(n_o+n_i)^{-1}]\ \mathcal{E}\{\tilde{\omega}^{-1}|\mathcal{D}\}\right]$, for all $t>n_i$.

To see these results in a more familiar context, consider the case of *one* risk, in which:

$$Y_o = y_1 \quad ; \quad Y_{oo} = y_{11} \quad ; \quad Z = \bar{z} = z_1 \quad ; \quad N = n_1 \quad . \tag{5.10}$$

and the exponent factor in (5.4) becomes simply:

$$B(\mathcal{D}) = (y_{11} - y_1^2) + (1-z_1)(m - y_1)^2 \quad . \tag{5.11}$$

Then, the posterior-to-data variance of any future value of $\tilde{x}_{1t}$ can be predicted from the credibility formula:

$$\mathcal{V}\{\tilde{w}_{1t}|\mathcal{D}\} = \left[\frac{n_o+n_1+1}{2\alpha-2+n_1}\right]\left\{\left[\frac{2\alpha-2}{n_o+1}\right](1-z_1)v_o+z_1(1-z_1)(m-y_1)^2+z_1(y_{11}-y_1^2)\right\}. \tag{5.12}$$

where $v_o = \mathcal{V}\{\tilde{w}_{1t}\} = (1+n_o^{-1})\ \mathcal{E}\{\tilde{\omega}^{-1}\}$. If we take the "natural" choice of $\alpha = (n_o+3)/2$ (see remarks in Jewell & Schnieper [1985]), we obtain the well-known natural-conjugate results for the one-dimensional normal with unknown mean and variance. In other words, the results of this section can be viewed as the generalization of this simple individual risk model when collateral data is available.

6. A General Hierarchical Model with All Variances Linked

We are now ready to tackle a general heteroscedastical hierarchical model. The success of the non-hierarchical model of the last section was due primarily to assumption (5.1), which removes the dependency of the z_i upon any random quantity. This suggests that an appropriate assumption in the hierarchical case would be to link all three precisions together, by assuming:

$$\tilde{\tau} = n_o\tilde{\omega}\ ;\ \tilde{\eta} = r_o\tilde{\tau} = r_o n_o\tilde{\omega}\ . \tag{6.1}$$

for appropriate positive n_o and r_o. This is equivalent to saying that . have very tight knowledge about how the total variance is split up among the three levels of the model, but the value of the total variance is unknown. This assumption not only simplifies the various credibility factors:

$$(6.2)\quad z_i = n_i/(n_o+n_i)\ ;\ z_o = Z/(r_o+Z)\ ;\ \bar{z} = \Sigma n_i z_i/N\ ;\ z_c = (1+\frac{r_o}{r})z_o\ .$$

but also makes the (f_i), f_o, Y_o, and f_c *independent* of $\tilde{\omega}$! Thus, the unconditioning on ω required in the general formulae of Sections 2 and 3 reduces to simple expectations over $(\tilde{\omega}|\mathcal{D})$, with the possibility of using the natural conjugate Gamma prior of (5.5)! For completeness, we now display all of the final results, using notation and arguments developed previously.

6.1 Posterior Parameter Results

If $p(\omega)$ is $Ga(\alpha,\beta)$, then from (2.11c) we see that $p(\omega|\mathcal{D})$ is $Ga(\alpha',\beta')$, with updated parameters:

$$\alpha' = \alpha + N/2 \quad ; \quad \beta' = \beta + NB(\mathcal{D})/2 \quad ;$$

$$(6.3)\qquad B(\mathcal{D}) = [Y_{oo}-Y_o^2] + (1-\bar{z})(1-z_o)[m-Y_o]^2 + N^{-1}\Sigma n_i z_i[Y_o^2-y_i^2]\ .$$

similar to (5.4). The most important moment formula for our purposes is $\mathcal{E}\{\tilde{\omega}^{-1}|\mathcal{D}\} = \alpha'/(\beta'-1)$, which we note can be written:

$$(6.4)\qquad \mathcal{E}\{\tilde{\omega}^{-1}|\mathcal{D}\} = (1-z_*)\ E\{\tilde{\omega}^{-1}\} + z_*\ B(\mathcal{D}),$$

where we have eliminated β in favor of the prior mean variance and have defined a *variance credibility factor*:

$$(6.5)\qquad z_* = N/(N+2(\alpha-1))\quad .$$

From (2.6) and (2.9), we find the posterior-to-data moments of the portfolio mean to be:

(6.6) $$\mathcal{E}\{\tilde{\varphi}|\mathcal{D}\} = f_o(\mathcal{D}) = (1 - z_o)\, m + z_o\, Y_o \quad ;$$

(6.7) $$\mathcal{V}\{\tilde{\varphi}|\mathcal{D}\} = \left[\frac{1 - z_o}{r_o n_o}\right] \mathcal{E}\{\tilde{\omega}^{-1}|\mathcal{D}\} \quad .$$

In fact, because of the position of ω in the simplified version of $p(\varphi|\Omega,\mathcal{D})$, we see that $p(\varphi|\mathcal{D})$ must be a one-dimensional Student-t density, with $2\alpha' = 2\alpha + N$ degrees of freedom, and the above mean and variance.

Progressing to the individual risk means, we use the results of Section 3 to find the individual mean forecasts as before:

(6.8) $$\mathcal{E}\{\tilde{\mu}_i|\mathcal{D}\} = f_i(\mathcal{D}) = (1 - z_i)\, f_o(\mathcal{D}) + z_i\, y_i \quad ,$$

and the new covariance structure:

(6.9) $$\mathcal{C}\{\tilde{\mu}_i;\tilde{\mu}_j|\mathcal{D}\} = \sigma_{ij}(\mathcal{D}) =$$

$$\left[\frac{(1-z_i)}{n_o}\,\delta_{ij} + \frac{(1-z_o)(1-z_i)(1-z_j)}{r_o n_o}\right] \mathcal{E}\{\tilde{\omega}^{-1}|\mathcal{D}\} \quad .$$

where δ_{ij} is the Kronecker delta-function, and the data values enter through the use of (6.3)-(6.5). Again, because of the normal-gamma distributional assumption, one can argue that $p(\underline{\mu}|\mathcal{D})$ is a r-dimensional Student-t density, with $2\alpha + N$ degrees of freedom.

6.2 Posterior Predictive Results

Passing to the forecasts of future values, (3.4) (3.5) give us simply:

$$(6.10)\qquad \mathcal{E}\{\tilde{w}_{it}|D\} = f_i(D) = (1 - z_i)\, m + z_i\, y_i \quad .$$

$$(6.11)\qquad \mathcal{C}\{\tilde{w}_{it};\tilde{w}_{ju}|\mathcal{D}\} = \delta_{ij}\ \delta_{ju}\ \mathcal{E}\{\tilde{\omega}^{-1}|\mathcal{D}\} + \sigma_{ij}(\mathcal{D}) \quad .$$

From these, various multivariate Student-t densities can be generated "over future times" for a single risk, "over risks" at one future epoch, or for various mixtures of risks and epochs.

Finally, the average future risk for the entire cohort (3.7)(3.8) has moments:

$$(6.12)\qquad \mathcal{E}\{\tilde{w}_o|\mathcal{D}\} = f_c(\mathcal{D}) = (1 - z_c)\, m + z_c\, Y_o \quad ;$$

$$(6.13)\quad \mathcal{V}\{\tilde{w}_o|\mathcal{D}\} = \left\{\frac{1}{r}\left[1 + \frac{1}{n_o}(1 - \frac{Z}{r})\right] + \left[\frac{1-z_o}{r_o n_o}\right]\left[1 - \frac{Z}{r}\right]^2\right\} \mathcal{E}\{\tilde{\omega}^{-1}|\mathcal{D}\} \quad .$$

It follows also from arguments similar to those above that $p(w_o|\mathcal{D})$ will be a Student-t density.

Any of the final (co)variance formulae can be put into a credibility-like form. For example, the formula corresponding to (5.12) for the total variance of a future observation of risk #1 can be written as:

$$(6.14)\qquad \mathcal{V}\{\tilde{w}_{1t}|\mathcal{D}\} = \left[\frac{r_o n_o + (1-z_1)r_o + (1-z_o)(1-z_1)^2}{r_o n_o + r_o + 1}\right] \times$$

$$\left\{(1-z_*)\mathcal{V}\{\tilde{w}_{1t}\} + z_* \left[\frac{r_o n_o + r_o + 1}{r_o n_o}\right]\left[(Y_{oo} - Y_o^2) + (1-\bar{z})(1-z_o)(m-Y_o^2) + N^{-1}\ \Sigma n_i z_i (Y_o^2 - y_i^2)\right]\right\} \quad .$$

7. Similar Data Lengths

Notation simplifies dramatically if all data record lengths are the same, i.e., $n_i = n$ $(i=1,2,\dots,r)$. Then:

(7.1) $N=nr$; $z_i=z=\bar{z}=n/(n_o+n)$; $Z=rz$; $z_o=rz/(r_o+rz)$; $z_*=nr/(nr+2(\alpha-1))$,

and the overall statistics Y_o and Y_{oo} can be replaced by the simpler versions:

$$(7.2) \qquad y_o = \Sigma y_i/r = \Sigma\Sigma x_{it}/nr \ ; \quad y_{oo} = \Sigma y_{ii}/r = \Sigma\Sigma x_{it}^{\ 2}/nr \ .$$

Then (6.3) simplifies to:

$$(7.3) \qquad B(\mathfrak{D}) = y_{oo} - (1-z)\, y_o^2 - z(\Sigma y_i^2/r) + (1-z)(1-z_o)(m-y_o)^2 \ .$$

For approximations and asymptotic studies, it is useful to eliminate y_o^2 and $\Sigma y_i^2/r$ in favor of the statistics:

$$(7.4) \qquad y_{o;o} = \Sigma y_{i;i}/r \quad ; \quad y_{i;i} = \left[\frac{2}{n(n-1)}\right]\sum_{t<u}\sum x_{it}\, x_{iu} \quad ;$$

$$y_{**} = \left[\frac{2}{r(r-1)}\right] \sum_{i<j}\sum y_i y_j \quad ,$$

from which one can show that:

$$(7.5) \qquad y_o^2 = \frac{1}{r}\left(\frac{1}{r}\, \Sigma y_i^2\right) + \left(\frac{r-1}{r}\right) y_{**} \quad ; \quad \frac{1}{r}\Sigma y_i^2 = \frac{1}{n}\, y_{oo} + \left(\frac{n-1}{n}\right) y_{o;o} \ .$$

$B(\mathfrak{D})$ and the variances above can be manipulated into various useful forms. For instance, $B(\mathfrak{D})$ can be expressed solely in terms of the differences $(y_{oo}-y_{o:o})$ and $(y_{o:o}-y_{**})$, the various credibility factors, and the design parameters, n and r.

8. Single Unknown Variance

To illustrate the difficulties caused by more general precision structures, we consider the model analyzed by Berger [1985, 4.6], in which it is assumed (in our notation) that $\omega = f^{-1}$ and $\eta = h^{-1}$ are known with certainty, so that the unknown paramters are $(\tilde{\underline{\mu}},\tilde{\varphi};\tilde{\gamma})$. For simplicity, we assume all data lengths are equal to n (Berger has n = 1).

Thus there are only two conditional credibility factors to consider:

$$(8.1) \qquad z_1(\gamma) = \frac{n}{f\gamma+n} \quad ; \quad z_o(\gamma) = \frac{(rnh)\gamma}{(rnh+f)\gamma+n} \quad ;$$

plus the individual statistics (y_i, y_{ii}) and the simplified portfolio statistics (y_o, y_{oo}). Other factors in (2.12) are simply: $N = nr$, $\bar{z}(\gamma) = z_1(\gamma)$, and $Z(\gamma) = rz_1(\gamma)$.

The posterior precision density (2.10) now becomes a one-dimensional formula that can be "simplified" to reveal the structure on γ:

$$(8.2) \qquad p(\gamma|\mathfrak{D}) \quad p(\gamma) \left[\frac{f\gamma}{f\gamma+n}\right]^{r/2} \left[\frac{f\gamma+n}{(rnh+f)\gamma+n}\right]^{1/2} \exp\,(-C(\gamma)/2) \;;$$

$$(8.3) \qquad C(\gamma) = (rnf^{-1}) \left[\left[\frac{n}{f\gamma+n}\right] \left[(y_o^2 - \Sigma y_i^2/r) + \left[\frac{f\gamma}{(rnh+f)\gamma+n}\right] (m-y_o)^2\right]\right].$$

It is clear that no analytical prior will lead to a tractable posterior density, so that numerical methods will have to be used in this case. However, examination of the results of Section 3 shows that the essential work in finding the various posterior and predictive first and second moments will be in taking expectations of various powers of the two credibility factors, plus finding $\mathcal{E}\{\tilde{\gamma}^{-1}|\mathcal{D})$. The situation is similar if either $\tilde{\omega}$ or $\tilde{\eta}$ is an unknown variance, except that, in the latter case, only z_o will be modified by the data.

9. Interpretation of Results

As previously mentioned, the various credibility results (the formulae with notation $f(\mathcal{D})$) show that predictive means can be represented as linear convex combinations of a prior mean with a classical estimator for that mean. With special priors, this may even be true for an updated variance (6.4). As the number of data points increases, the credibility factor increases, so that we place more "credibility" in the classical estimator, until (in the simplest cases), the Bayes estimator is essentially the sampling-theory estimator. The rate at which this "learning" occurs depends upon a credibility "time constant" that describes how our prior uncertainty is split between observational variation and prior uncertainty in the mean parameter.

In the hierarchical model with known variances, (3.2) and (2.6) reveal that the individual mean forecasts, $f_i(\mathcal{D})$, are linear combinations between the classical estimators, y_i, and another credibility

forecast for the portfolio-level mean, $\tilde{\varphi}$. $f_o(\mathfrak{D})$ itself mixes a "universal" mean, m, with a portfolio-wide statistic, Y_o, according to a cohort learning curve, z_o, with a time constant that is the ratio of higher level variances. Note that the individual data (x_{it}) from risk i all has the same unity weight over time in calculating the individual-level statistic, y_i, but that these then have differing weights, z_i, over risks in calculating the portfolio statistic, Y_o. Further, the z_i increase with increasing numbers of samples, n_i, but z_o increases only with increasing sum Σz_j over the number of risks, r! It can be shown that this appealing layer-by-layer compounding of credibility forecasts that depend only upon local sources of variation and lower-level credibility factors repeats itself in more general hierarchical models with many layers (Bühlmann & Jewell [1987]). An important practical remark is that, while the z_i approach unity with increasing n_i, this is *not* true at the portfolio level, since z_o remains less than unity for finite r. In other words, because our finite portfolio may represent a *biased* selection of possible $(\tilde{\theta}_i)$, Y_o is not "fully credible" for $\tilde{\varphi}$, even with very large numbers of data points per risk.

In fact, it was examination of the credibility factors that revealed the essential simplification in the generalized hierarchical model (2.2). In order to keep the intuitive and appealing linear formulae for the predictive means, we *must* constrain the ratios of adjacent-level precisions so that the credibility time constants, n_o and r_o, will remain constant, so that only one precision is unknown! This has the very useful

byproduct of making the exponent A in (2.10)(2.11) linear in that precision, so that the Gamma is a natural-conjugate prior density. The notation n_o and r_o suggests that the prior parameter densities behave *as if* prior information were equivalent to r_o risks each having n_o data samples with common value, m. Similar constraints are necessary in more complicated hierarchical models to keep the predictive means in credibility form; full details for the Lindley & Smith [1974] linear model will be in a forthcoming paper.

There are many other combinations of model and prior densities that have simple credibility forecasts using individual data. However, the author has been unable to extend these models to the hierarchical case and still retain the linear structure for a portfolio-level forecast using collateral data. I would be interested in hearing from others who have looked further into special structures for first- and second-moment forecasts in hierarchical models.

REFERENCES

Berger, J.O. (1985). *Statistical Decision Theory and Bayesian Analysis*. New York: Springer-Verlag 2nd ed.

Bühlmann, H. (1967). Experience rating and credibility, *ASTIN Bulletin* 4, 199-207.

Bühlmann, H. & Jewell, W.S. (1987). Hierarchical credibility revisited. *Bulletin Association of Swiss Actuaries* 1, 35-54.

Jewell, W.S. (1974). Credible means are exact Bayesian for simple exponential families, *ASTIN Bulletin* 8, 77-90.

Jewell, W.S. (1975a). Regularity conditions for exact credibility, *ASTIN Bulletin* 8, 336-341.

Jewell, W.S. (1975b). The use of collateral data in credibility theory: a hierarchical model, *Giornale dell'Istituto Italiano degli Attuari* 38, 1-16.

Jewell, W.S. & Schnieper, R. (1985). Credibility approximations for Bayesian prediction of second moments, *ASTIN Bulletin* 15, 103-121.

Jewell, W.S. (1987). A general framework for second moment multi-dimensional credibility approximations, *O.R. Center, U. California, Berkeley*, in preparation.

Jewell, W.S. & Bühlmann, H. (1987). Credibility approximations for second moments in hierarchical models, *O.R. Center, U. California, Berkeley*, in preparation.

Lindley, D.V. & Smith, A.F.M. (1972). Bayes estimates for the linear model, *Journal Royal Statistical Society B*, 1-41.

END

10-87

DTIC